[This Page Intentionally Left Blank]

8.0 3T3 AND NHK NRU TEST METHOD DATA QUALITY

This section of the BRD presents the extent of adherence to national and international GLP guidelines during for generation of the NICEATM/ECVAM validation study data. Data quality is described along with any deviations from the guidelines and the impact of any noncompliance. Statistical results are provided to show comparison of data generation, collection, and reporting of the two GLP adherent cytotoxicity testing laboratories and the one non-GLP adherent cytotoxicity testing laboratory as well as the GLP laboratory that distributed the reference substances and performed solubility studies. Discussions of various quality assurance aspects of the study are included.

8.1 Adherence to Good Laboratory Practice Guidelines

8.1.1 Guidelines Followed for *In Vitro* NRU Cytotoxicity Testing

Good Laboratory Practices

The SOW provided the following definition of U.S. Regulatory agency GLPs to each laboratory:

> *"Regulations governing the conduct, procedures, and operations of toxicology laboratories; regulations to assure the quality and integrity of the data and to address such matters as organization and personnel, facilities, equipment, facility operations, test and control articles, and validation study protocol, and conduct (U.S. Food and Drug Administration, Title 21 CFR Part 58; U.S. Environmental Protection Agency, Title 40 CFR Part 160)."*

IIVS, ECBC, and BioReliance performed testing under all GLP guidelines. The details of GLP compliance and training are addressed in **Section 11**.

Spirit of GLP

The SMT determined a definition for "spirit of GLP" and provided the following verbiage to the laboratories:

"Laboratories that are non GLP-compliant shall adhere to GLP principles and other method parameters as put forth in this Statement of Work and the Test Method Protocols (provided by NIEHS/NICEATM); documentation and accountability shall be equal to GLP requirements; laboratories must make assurances that they are equal in performance criteria and that there is parity amongst the laboratories."

FAL performed testing in the "spirit of GLP" (see **Section 11.2.2**) by following the international GLP standards referenced in the ECVAM Workshop 37 Report (Cooper-Hannan 1999) and the OECD Principles of GLP (OECD 1998). The laboratory did not have data and test method procedures reviewed by an independent quality assurance (QA) auditor. At a minimum, the SOW directed FAL to routinely document the following laboratory tasks (e.g., equipment monitoring) and record keeping (see **Table 8-1**) and to archive the documents. The FAL laboratory already had most of the following procedures and guidelines in place for routine laboratory procedures before initiation of this study. The various general laboratory-related activities were documented in workbooks and logbooks and the information was made available to the SMT.

Table 8-1 SMT-Recommended Documentation for FAL Laboratory

Daily	Per Use	Periodic
Temperatures Laboratory, incubators, water baths, refrigerators, freezers	Cryogenic Storage Unit Liquid N_2 volume	Laboratory Supplies[1] Lot numbers and expiration dates for stock media formulations and components, NRU reagents, tissue culture plasticware
Humidity/CO_2 Cell culture incubators	Equipment Calibration Balances, pH meters, and cell counters	Cells Quantity and cryogenic storage conditions for 3T3 and NHK cells
Visual Observations Cell Culture Growth	Reagents Lot numbers and expiration dates of medium/supplements	Equipment Calibration Incubators, laminar flow hoods, autoclaves, micropipettors, spectrophotometer plate readers, computers (software)

[1]Periodic documentation for laboratory supplies occurs when supplies are purchased and received in the laboratory

Good Cell Culture Practices (GCCP)

The SMT provided guidance in the SOW for implementing GLPs in a cell culture laboratory environment. The initial assumption by the SMT was that each laboratory had the basic cell culture skills and knowledge (e.g., as described in Freshney 2000) to perform the NRU cytotoxicity test methods in a reliable manner. Reviews of historical documents and scientific and professional exchanges with the laboratory personnel assured the SMT that each laboratory had demonstrated, through previous validation studies and other scientific endeavors, that personnel were capable of providing quality scientific data through the use of good cell culture practices. A comparison of the SOW and the *in vitro* NRU cytotoxicity protocols to the ECVAM Good Cell Culture Practices (GCCP) Reports (Hartung 2002; Coecke et al. 2005) and the OECD document on GLPs and *in vitro* studies (OECD 2004a) showed that the guidelines in place for the NICEATM/ECVAM study were harmonious with the ECVAM and OECD guidelines.

8.1.2 Quality Assurance (QA) for *In Vitro* NRU Cytotoxicity Test Data

Coded Reference Substances

BioReliance acquired 73 high purity chemicals (72 reference substances and one positive control chemical at 99% or greater purity when economically feasible) from reputable commercial sources according to the SOW provided by the SMT (see **Appendix G**). Seven reference substances were less than 99% pure (three less than 98% pure; lactic acid had the lowest purity [89%]). The substances were coded with unique identification numbers and provided to the testing laboratories in a blinded fashion. Preparation of substances for distribution was performed under GLP guidelines. **Section 3.6** provides detailed information concerning acquisition and distribution of reference substances.

Solubility Testing and Data Review

All laboratories performed solubility tests on all reference substances using the solvents and procedures specified by the protocols provided by the SMT and submitted solubility data as hard copy printouts and electronic worksheets. The laboratories also maintained solubility data in their workbooks. The Study Directors reviewed all laboratory procedures and all data produced at their respective laboratories. The QA designee reviewed all data in the GLP-

adherent laboratories. The SMT Project Coordinators served as informal QA reviewers for the FAL (i.e., reviewed all raw data sheets). Detection of errors and omissions were reported to FAL and corrections were requested. The SMT reviewed all solubility data and all NRU assay data produced by all laboratories for this study.

The SMT reviews of submitted data in Phases Ia and Ib revealed that even after data review by the Study Directors, data files contained an unacceptable high frequency of errors (see **Section 2.6.3**). The laboratories were alerted to the problem and personnel from all the laboratories attended a weeklong training session to enhance harmonization among the laboratories. After the training, errors were still found in data files submitted for Phase III, albeit less frequently; such errors generally occurred due to the rapid submission of data files to the SMT shortly after the conclusion of each test. The formal QA review of the files occurred later in each phase of the study.

Errors included typographical mistakes, transcriptional and data entry errors in the Microsoft® EXCEL® and the GraphPad PRISM® 3.0 templates, and incorrect labeling of files. The SMT reviewed every electronic file and hard copy printout throughout the study and alerted the Study Directors when errors were found. All data files were checked for consistency within the documents and for compliance with the protocols. The SMT also documented errors on the hard copy printouts as handwritten notations and included these notations in the electronic data summary files compiled for data management. Files that were revised and/or corrected by the Study Director were resubmitted to the SMT and noted as corrected files.

In Vitro NRU Cytotoxicity Test Tallies

Periodically, the laboratories received individualized test tallies from NICEATM that detailed:

- the number of range finder tests performed
- the number of definitive tests performed and the pass/fail status of each test
- the number of positive control assays performed and the pass/fail status of each test

- the number of acceptable tests completed per the SMT and protocol requirements
- the status of test completion for each substance (i.e., whether one range finder test and three acceptable definitive tests had been completed for the substance)

The laboratories compared the NICEATM tallies to their own records to verify consistency and accuracy. Discrepancies were resolved through direct communication between the Study Director and the SMT.

8.1.3 Guidelines Followed for *In Vivo* Rodent Oral LD_{50} Data Collection

The *in vitro* NRU cytotoxicity test methods are proposed as methods to predict starting doses for acute oral lethality *in vivo* (specifically, rat) assays and not as replacement tests for an *in vivo* reference method. No *in vivo* tests were performed for this validation study. All *in vivo* data (i.e., rodent [rat and mouse] LD_{50} values) were collected by NICEATM through reviews of the literature. All data and pertinent information were gathered and stored in a spreadsheet database.

Rodent Acute Oral LD_{50} Values Used in the Registry of Cytotoxicity (RC)

The RC rodent (rat and mouse) acute oral LD_{50} values came largely from the 1983/84 RTECS® database (compiled by NIOSH). The RC is a database of acute oral LD_{50} values for rats and mice obtained from RTECS® and IC_{50} values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for chemicals with known molecular weights (Halle 1998). Collection and reporting methods used for generating the data were not a part of any data collection hierarchy employed by the NIOSH. The data in the RTECS® database were not evaluated for quality and accuracy by NIOSH. Many sources of the values come from secondary references with no citation for the original report. GLP guidelines for acute oral toxicity testing were not part of any criteria for determining acceptable data for the database. The only criterion the NIOSH used for reporting acute oral toxicity data in RTECS® was that the LD_{50} value was the most toxic LD_{50} value for a chemical that could be found in the literature.

Rodent Acute Oral LD_{50} Values Collected by NICEATM

One critical aspect of the study design was the establishment of a rat acute oral LD_{50} reference value for each of the 72 reference substances (see **Section 4**). These reference values were used to evaluate the extent to which the two *in vitro* test methods can predict rat acute oral LD_{50} values. Primary rat acute oral LD_{50} studies were located through searching electronic databases, published literature, and secondary references. Rat data were not available for three of the reference substances and, for these, mouse acute oral LD_{50} values were collected. Very little data collected from the literature were produced under GLP guidelines; in fact, only seven of the 455 LD_{50} values collected were obtained under GLP conditions.

8.2 Results of Data Quality Audits

The QA unit or designee of each GLP laboratory provided a systematic and critical comparison of the data provided in the study report to the raw data in the laboratory records. The SOW provided to each laboratory contained the following guidance on QA statements:

> *"The Final Reports for all phases of the Validation Study shall be audited by the Quality Assurance unit of the Testing Facility for GLP compliance and a QA Statement shall be provided by the Testing Facility. Each Final Report shall identify: 1) the phases and data inspected, 2) dates of inspection, and 3) dates findings were reported to the Study Director and Testing Facility management. The QA Statement shall identify whether the methods and results described in the Final Report accurately reflect the raw data produced during the Validation Study."*

8.2.1 QA Statements

The QA statements from the GLP-compliant laboratories noted the QA reviews of:

- protocols
- laboratory standard operating procedures (SOPs)
- laboratory operations
- 3T3 and NHK NRU experiment data
- final report

The QA statements report that the test methods described in the protocols are the methods that the laboratory personnel used and that the data reported to the SMT is an accurate reflection of the raw data obtained by the laboratory. See **Section 8.2.2** for information about the QA statements for the non-GLP laboratory.

8.2.2 QA Statements from the Laboratories

BioReliance QA Statements

The Study Director/Laboratory Director provided the following statement in all of the final reports from BioReliance:

> *"The solubility studies, acquisition, preparation, and distribution of the test chemicals were conducted in compliance with GLP. Although not audited (per SOW), the work described in this report for Phase X (i.e., Ia, Ib, and II) fully and accurately reflects to the best of my knowledge the raw data generated in the study."*

FAL QA Statements

The Study Director for the FAL laboratory performed the final review of all data and reports before sending to the SMT and provided two statements in the final reports (provided to the SMT).

- *"The laboratory worked under the principles of GLP whilst not being a GLP-compliant laboratory."*
- *"The report accurately reflects the work undertaken and the results obtained at the FRAME Alternatives Laboratory."*

Since the SMT performed QA reviews of the FAL as an informal reviewer, formal QA statements were not provided to FAL.

ECBC QA Statements

The QA statements reported what particular study phase and which laboratory procedures were examined for compliance with GLP guidelines. In addition, the statement reiterated that the scope of work, associated protocols, and quality control acceptance criteria were

updated/changed during the study which made it more difficult to assess the procedures and data for conformance to the protocols. However, during the review of SOPs and the observance of operations, the requirements and intent of GLP guidelines were continually assessed. The QA reviews found the ECBC protocols to be in compliance with the NICEATM/ECVAM study protocols. The phases of the studies inspected by the QA designee were as follows:

- review of protocols and laboratory SOPs
- review of waste handling
- review of laboratory operations
- certification of new personnel
- review of data
- review of the final report for each phase

The QA designee also observed preparation of reference substances, 96-well plate configuration, application of reference substance, annotation to the workbook, and appropriate sterile technique while performing the testing. The number of inspections of laboratory operations were reduced in the latter phases of the validation study since the same personnel conducted the testing throughout the entire study.

<u>ECBC Review Dates of Various Aspects of the Study</u>

- Phase Ia: July 2002 through May 2003
- Phase Ib: July 2002 through January 2003
- Phase II: May 2003 through February 2004
- Phase III: November 2003 through March 2005

IIVS QA Statements

Because the IIVS QA unit is small, it carried out reviews in phases. The IIVS QA Statement reads:

> "This study has been divided into a series of in-process phases. Using a random sampling approach, Quality Assurance monitors each of these phases over a series of studies. Procedures, documentation, equipment records, etc., are examined to assure

that the study is performed in accordance with the U.S. FDA Good Laboratory Practice regulations (21 CFR 58), the U.S. EPA GLP Standards (40 CFR 792 and 40 CFR 160) and the OECD Principles of Good Laboratory Practice and to assure that the study is conducted according to the protocol and relevant Standard Operating Procedures."

The phases of the studies inspected by the QA designee were as follows:

- protocol and initial paperwork
- reading of the plates (definitive assay)
- dilution of the test articles (definitive assay)
- termination of treatment and addition of the NR dye (definitive assay)
- cell concentration determination and seeding of the plates (third definitive)
- termination of treatment and addition of the NR dye
- washing the cells
- treatment of the cells
- draft report and data
- final report

<u>IIVS Review Dates of Various Aspects of the Study</u>

- Phase Ia: August 2002 Final Report Review: October 2005
- Phase Ib: January 2003 Final Report Review: October 2005
- Phase II: July-August 2003 Final Report Review: October 2005
- Phase III: January-November 2004 Final Report Review: October 2005

Other QA Information

Data generated by the laboratories and reviewed by their respective Study Directors were provided directly to the SMT. Often, the data were provided electronically within days of the end of testing. The SMT was very active as a secondary QA reviewer concerning all information provided by the Study Directors. If the SMT found discrepancies, then the Project Coordinators corresponded with the appropriate Study Director to rectify the mistake.

The Study Director made corrections/adjustments to any discrepancies in data reporting and presented any changes to the SMT. The SMT did not initiate any external data quality audits.

The quality of the reference substances was assured in the form of certificates of analysis provided by the chemical manufacturer to BioReliance at the time of purchase. The SMT and the laboratories obtained certificates of analysis from CAMBREX specifically for Clonetics® NHK culture medium and supplements. In addition, the SMT obtained quality control data directly from CAMBREX technical departments for determining the NHK medium's ability to support keratinocyte growth.

8.3 Impact of Deviations from GLPs/Non-compliance

Several error rates were determined by the SMT in regard to documentation, testing methods, and data manipulation by the laboratories. Many errors (particularly in Phases Ia and Ib) were minor mistakes (e.g., typographical, mislabeling) and did not affect the quality of the data.

8.3.1 Laboratory Error Rates

During Phases Ia and Ib, the SMT was concerned about the number of errors in documentation and testing methods and compiled the number of detected errors from each laboratory. The types of errors noted and compiled included errors in documentation (e.g., reference substance identification did not match on all associated data sheets, IC_{20} and IC_{80} values were switched in the EXCEL® template, a test acceptance criterion flag in data sheet was incorrect, etc.) and in testing (e.g., wrong dilution scheme was used for the PC, wrong SLS IC_{50} was used as the PC IC_{50}, etc.). Error rates were compiled as number of tests with errors per total number of tests. As shown in **Table 2-3**, FAL had the highest error rates: 93% for the 3T3 assay and 41% for the NHK assay. The highest error rates of the other laboratories were 10% for the 3T3 assay and 23% for the NHK assay (both ECBC).

There were very few errors detected in the Phase III data files. The SMT did not compile typographical and transcriptional errors but reported the errors directly to the appropriate

Study Director so that the data sheets could be immediately rectified. The SMT did not detect errors in the raw optical density data from the 96-well plates provided in each data file. The laboratories and the SMT corrected any typographical and transcriptional errors (e.g., incorrect logIC_{50} value entered) in the EXCEL® templates. The template formulas calculated the correct values for the statistical analyses and the quality of the data was not compromised.

For Phase III, assessment of error rates was performed specifically for Phase III for one particular clerical error – the transfer of statistical results (e.g., ICx values) from the GraphPad PRISM® 3.0 template to the Microsoft® EXCEL® template. It was often necessary for the SMT to revise the Microsoft® EXCEL® data files provided by the laboratories because the incorrect values had been transferred to the template. The SMT revised files (using the data in the PRISM® 3.0 template) due to this error and reports as follows as the number of errors/total number of definitive tests:

Table 8-2 Error Rates

Laboratory	Number of Errors Detected[1]	Number of Definitive Tests	Percentage of Tests with Detected Errors
ECBC	49	402	12
FAL	171	513	33
IIVS	25	419	6

[1]Clerical error – transfer of statistical results from PRISM® to EXCEL®

8.3.2 Test Failure Rates for Definitive Tests and PC Tests

Table 8-3 illustrates the test failure rates experienced for Phase III of the validation study. Approximately 25% of all 3T3 definitive tests and 18% of all NHK definitive tests failed (i.e., did not meet test acceptance criteria). If a definitive test (see **Section 2.2.2** for the definition of a definitive test) failed, then the laboratory repeated the test and attempted to reach the goal of three acceptable definitive tests for each reference substance and each cell type (see **Section 2.5** for criteria for repeating tests). PC failure occurred 0 – 18% of the time with an overall average failure rate of 8% combined for both assays. FAL had the highest individual laboratory test failure rates for 3T3 definitive tests (30%), NHK definitive tests (32%), and NHK PC tests (18%). ECBC had the highest failure rate for 3T3 PC tests (11%).

Phase III guidelines called for each laboratory to provide three acceptable definitive tests for each substance for both cell types (3 x 60 x 2 = 360 definitive tests). PC tests were run concurrently with the definitive tests and generally more than one reference substance was tested in conjunction with one PC test plate. Due to test failures, each laboratory performed additional testing to attempt to obtain the three acceptable definitive tests requested for each substance.

Table 8-3 Definitive Test and Positive Control (PC) Test Failure Rates

Test Type	3T3 NRU Test Method				NHK NRU Test Method				Total
	ECBC	FAL	IIVS	Total	ECBC	FAL	IIVS	Total	
Definitive Tests - Acceptable	169	177	176	522	173	175	174	522	1044
Definitive Tests - Total	215	257	225	697	187	256	194	637	1334
% Definitive Tests Failed	21	30	22	25	8	32	10	18	22
PC Tests - Acceptable	66	40	16	122	58	37	20	115	237
PC Tests - Total	74	42	17	133	59	45	20	124	257
% PC Tests Failed	11	5	6	8	2	18	0	7	8
Definitive Tests Failed Only Because PC Tests Failed	14	6	14	34	0	22	0	22	56
% Definitive Tests Failed Only Because PC Tests Failed	7	2	6	5	0	9	0	4	4

Table 8-4 illustrates the success rates of the testing for each laboratory and for the combined laboratories.

Table 8-4 Definitive Test and PC Test Success Rates for 3T3 and NHK NRU Test Methods (Combined Total Tests)

Test Type	ECBC	FAL	IIVS	Total
Acceptable Definitive Tests/ Total Definitive Tests	342/402	352/513	350/419	1044/1334
% Acceptable Definitive Tests	85%	69%	84%	78%
Acceptable PC Tests/Total PC Tests	124/133	77/87	36/37	237/257
% Acceptable PC Tests	93%	89%	97%	92%

8.3.3 Intralaboratory Reproducibility

CV values for each reference substance were determined for each laboratory using the IC_{50} values from the acceptable definitive tests as described in **Section 5.3.1**. **Table 8-5** illustrates the average CV values for the substances tested in each of the phases and for the entire study.

Table 8-5 Coefficients of Variation

Cell Type	Labs	Phases I & II		Phase III		All Phases	
		Number of Reference Substances	Average % CV	Number of Reference Substances	Average % CV	Number of Reference Substances	Average % CV
3T3	ECBC	12	17	57	24	69	23
	FAL	11	28	55	33	66	33
	IIVS	11	20	56	22	68	21
NHK	ECBC	12	24	57	22	69	23
	FAL	12	31	57	45	69	42
	IIVS	12	14	58	14	70	14

8.3.4 Globally Harmonized System Toxicity Category Predictions

Predicted LD_{50} values were compared to the GHS *in vivo* acute oral toxicity categories to determine category match (i.e., accuracy) or toxicity underprediction or overprediction for the reference substances (see **Table 8-6**). Predicted LD_{50} values were determined for the reference substances by using the mean IC_{50} values from the laboratories in the RC regression. The reference GHS *in vivo* acute oral toxicity category presented in **Table 8-6**

was the initial LD_{50} value used to select the substances (see **Table 3-1**). The laboratories were generally in agreement with each other in the predictions. Although FAL had the highest error rates and CV values, their predictions of GHS toxicity category using these NRU methods were consistent with the other laboratories. (See **Appendix J** for additional laboratory comparisons for the other *in vitro – in vivo* regressions evaluated in **Section 6**.)

Table 8-6 GHS Toxicity Category Predictions by Laboratory[1]

	Labs	Total Reference Substances	Category Match	Toxicity Overpredicted	Toxicity Underpredicted
3T3	ECBC	69	29%	41%	30%
	FAL	67	28%	43%	28%
	IIVS	69	28%	41%	32%
NHK	ECBC	69	28%	42%	30%
	FAL	69	28%	41%	32%
	IIVS	70	29%	40%	31%

[1]GHS-Globally Harmonized System categories of acute oral toxicity with LD_{50} in mg/kg (UN 2003). 3T3 and NHK NRU test method IC50 data (geometric mean of within laboratory replicates) used with the RC regression: $\log(LD_{50}\ \text{mmol/kg}) = 0.425 \times \log(IC_{50}\ \text{mM}) + 0.625$.

8.4 Availability of Laboratory Notebooks

All laboratories maintained laboratory notebooks patterned after a template provided by IIVS and provided copies of them to the SMT (archived at NICEATM) after each phase. The workbooks contained information from all aspects of testing including but not limited to:

- environmental conditions
- reagent identification
- preparation of 96-well plates
- preparation of reference substances
- treatment of cell cultures
- visual observations of cell cultures
- NRU assays
- data analysis

8.5 Summary

- Various determinations of test method and data collection errors consistently showed that FAL had the highest error level; however, the laboratory's GHS acute oral toxicity category predictions were comparable to the other laboratories' results. Data were not adversely affected by general transcriptional errors.
- The laboratories reported no significant deviations from the test method protocols and deviations that did occur during the testing phases were generally quickly acknowledged and addressed by the Study Directors. If a deviation occurred that would affect data (e.g., improper concentration of DMSO solvent), then that Study Director would reject the test, notify the SMT, and perform an additional test. Improper transfer of data to either the EXCEL® or PRISM® templates, which would affect the data, were recognized, documented, and rectified by the Study Director and/or the SMT.
- The SMT was diligent in reviewing all data sheets to ensure that data were not inadvertently attributed to the incorrect data summary files and that the correct data were used in all statistical analyses.

An electronic copy of all data for this validation study can be obtained upon request from NICEATM.

[This Page Intentionally Left Blank]